Université Paris Cité
UFR de Mathématiques et Informatique
45, rue des Saints-Pères, 75006, Paris.

M1 MMA 2023-2024
Optimisation

Feuille de TD n°1

Différentiabilité, gradient, hessienne.

Exercice 1 (Échauffement)
Soit $n, p \in \mathbb{N}^*$ et $f : \mathbb{R}^n \to \mathbb{R}^p$.

1. Que pouvez-vous dire de l'assertion suivante : f est différentiable sur $\mathbb{R}^n$ si et seulement si il existe $\varphi : \mathbb{R}^n \to \mathbb{R}$ tel que pour tout $x \in \mathbb{R}^n$, $h \in \mathbb{R}^n$, $f(x+h) = f(x) + \varphi(h) + o(h)$.
2. On suppose $n = p = 1$ et f dérivable sur $\mathbb{R}$. Montrer alors que f est différentiable sur $\mathbb{R}$ et exprimer la différentielle de f en tout point $x \in \mathbb{R}$ en fonction de f'.
   Quelle est la différentielle de la fonction cos sur $\mathbb{R}$?
3. Soit $f : \mathbb{R}^n \to \mathbb{R}^p$. L'application gradient ∇f est-elle bien définie ?
4. Exprimer le gradient de log en tout $t \in \mathbb{R}$.

Exercice 2 (Vers une interprétation géométrique du gradient)
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie pour tout $x \in \mathbb{R}^2$, par $f(x) = \|Ax\|_2^2$, où

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}.$$

1. Quelle est la régularité de la fonction f ?
2. Déterminer pour tout $x \in \mathbb{R}^2$, $h \in \mathbb{R}^2$, $\mathrm{d}f(x)(h)$, puis $\nabla f(x)$.
3. Déterminer pour tout $x \in \mathbb{R}^2$, $h, k \in \mathbb{R}^2$, $\mathrm{d}^2 f(x)(h,k)$, puis $\nabla^2 f(x)$.
4. a) Représenter graphiquement l'ensemble $L_1 = \{x \in \mathbb{R}^2 : f(x) = 1\}$, appelée ensemble de niveau 1 de f.
   b) Choisissez $x \in L_1$ et représenter le vecteur $\nabla f(x)$ en ce point. Que remarquez-vous ?

Exercice 3
On fixe $\varepsilon > 0$ et $n \in \mathbb{N}$ avec $n \geqslant 3$. On définit

$$\forall x \in \mathbb{R}^n, \quad J_\varepsilon(x) = \sum_{i=2}^{n-1} N_\varepsilon(x_{i+1} + x_{i-1} - 2x_i), \quad \text{où} \quad N_\varepsilon(t) = \sqrt{\varepsilon + t^2}.$$

1. Écrire le développement à l'ordre 1 de J_ε et en déduire que J_ε est différentiable, puis donner une expression de $\mathrm{d}J_\varepsilon(x)(h)$ pour tout $x, h \in \mathbb{R}^n$.
2. Montrer que J_ε peut s'écrire sous la forme

$$\forall x \in \mathbb{R}^n, \quad J_\varepsilon(x) = \sum_{i=2}^{n-1} N_\varepsilon(A_i x),$$

avec des matrices A_i, $i \in \{2, \dots, n-1\}$, que l'on explicitera. Réécrire la différentielle grâce aux A_i, et en déduire le gradient et la hessienne de J_ε en tout point $x \in \mathbb{R}^n$. *On pensera à justifier que J_ε est deux fois différentiable sur $\mathbb{R}^n$.*

3. Donner une expression de $\mathrm{d}^2 J_\varepsilon(x)(h,k)$ pour tout $x \in \mathbb{R}^n$ et $(h,k) \in \mathbb{R}^n \times \mathbb{R}^n$ faisant intervenir les A_i.

Exercice 4
Soit $N, M \in \mathbb{N}^*$. On considère la fonction suivante

$$J_\lambda : \alpha \in \mathbb{R}^N \mapsto \frac{1}{n}\sum_{i=1}^{n} \log(1 + e^{-y_i\langle \alpha, x_i\rangle}) + \lambda\|L\alpha\|^2,$$

où $\lambda > 0$, $L \in \mathcal{M}_{M,N}(\mathbb{R})$ et les x_i, y_i, pour tout $i \in \{1, \ldots, n\}$, sont des éléments fixés.

1. Donner la nature des éléments x_i, y_i, $i \in \{1, \ldots, n\}$. Sur quel espace est définie la norme euclidienne utilisée dans l'expression de J_λ ?
2. Donner une expression du gradient et de la hessienne de J_λ après avoir préalablement justifié qu'elle est deux fois continûment différentiable. *Indication : on pourra penser à introduire diverses notations comme dans l'Exercice 3 pour se faciliter la tâche.*

Exercice 5 (Formalisation de l'interprétation géométrique du gradient)
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. Pour $\lambda \in \mathbb{R}$, on note $L_\lambda = \{x \in \mathbb{R}^2 : f(x) = \lambda\}$ la ligne de niveau λ de f.

Soit $x_0 \in \mathbb{R}^2$ tel que $\nabla f(x_0) \neq 0$ et notons $\lambda_0 = f(x_0)$. On se donne $\gamma :]-\varepsilon, \varepsilon[\to \mathbb{R}^2$, pour un certain $\varepsilon > 0$, une fonction de classe $\mathcal{C}^1$ telle que $\gamma(0) = x_0$ et $\|\gamma'(0)\| = \|\nabla f(x_0)\|$.

1. Quelle condition doit-on avoir sur γ pour que $f \circ \gamma$ décroîsse le plus vite au voisinage de 0. En déduire que $-\nabla f(x_0)$ donne la direction de la plus forte pente de f en x_0.
2. On suppose maintenant que $f \circ \gamma$ est constante. Démontrer que $\nabla f(x_0)$ est orthogonale $\gamma'(0)$, i.e. à la tangente à la ligne de niveau L_{λ_0}.
3. On suppose que $\gamma'(0) = \nabla f(x_0)$. Montrer alors qu'il existe une fonction $\xi : \lambda \mapsto \xi(\lambda) \in \mathbb{R}^2$ définie sur un voisinage de λ_0, $\mathcal{C}^1$ et telle que $\xi(\lambda) \in L_\lambda$. En déduire un équivalent de $\|\xi(\lambda) - x_0\|$ lorsque $\lambda \to \lambda_0$. Interpréter.